MULTIRESOLUTION MRF-BASED TEXTURE SEGMENTATION USING THE WREATH PRODUCT TRANSFORM PHASE

Gagan Mirchandani and Xuling Luo

Department of Electrical and Computer Engineering
The University of Vermont
Burlington, Vermont 05405
{mirchand, xulinluo }@emba.uvm.edu

ABSTRACT

In this paper, we bring local phase and multiresolution analysis to the texture segmentation problem. A Markov random field characterization is still employed, except it is used to model *phase* correlations rather than *intensity* correlations. Since statistical characteristics of phase are typically quite different to those of intensity, there exists the potential for creating greater discrimination in its feature space. We apply the Wreath Product Transform and use the phase at higher scales to initiate the segmentation process. For textures defined by homogeneous regions of dominant local edges, we see that the new algorithm yields better segmentation than that obtained through conventional multiresolution algorithms based on lowpass data.

1. INTRODUCTION

The idea of using a Markov random field (MRF) to model texture, be it for intensity or phase, is an appealing one. With a MRF, each point is statistically dependent only on its neighbors. Hence a MRF model is well suited to account for spatial or phase dependencies and thus can be employed to impose constraints on the classification of nearby pixels. Segmentation is then treated as a statistical estimation problem. MRF models have been used to model segmentation labels [2, 4, 3]. Due to the equivalence of a MRF and Gibbs distribution [4, 3], the maximum *a posteriori* (MAP) estimate of the unknown labels based on the observations is obtained by the minimization of an energy function. Stochastic relaxation such as simulated annealing (SA) and deterministic relaxation schemes such as iterated conditional modes (ICM) [3] have been used for the minimization.

Allowing for interaction amongst neighboring pixels, results in minimization schemes for MAP estimation that carry a cost: a substantial increase in computation complexity. An alternative to the relaxation schemes is multiresolution analysis, which provides more efficient computation. Some authors have used multiresolution label fields. Others, [9, 12, 6] both multiresolution observation and multiresolution label fields. In either case, image segmentation is obtained through a coarse-to-fine process. After an initial estimate for the segmentation at a coarse scale, the estimate is refined with a MRF scheme. Results from the coarse scale are used as starting estimates for the next finer scale. When using multiresolution observation fields, an image pyramid is constructed using lowpass filters such as Gaussian [9], Haar [12], and Binomial [6].

The above framework often works adequately when the statistical properties of an image are generally distinct for different texture segments. Properties typically considered are gray level-based distributions. However, it is not uncommon for regions with the same texture to have completely different distributions or for different textures to have similar distributions. We see in [13] examples of both categories: four regions of the same brittle fracture structure have completely different distributions, while two different fractile structures (brittle and ductile) have very similar distributions. Accordingly, we investigate the role of phase-based distribution for segmentation. Many situations exist, such as in SAR interferometry [5], where a phase characterization allows the determination of topographic elevation and small ground deformation. Hence, a multiresolution local phase-based image segmentation algorithm is proposed in this paper. The approach consists of five steps: (1) Use the Wreath Product Transform (WPT) [7] to construct the multiresolution observation fields and use the multiscale phase at scale 1 as the texture features. (2) Estimate the Gaussian distribution parameters for each of the predetermined number of classes. (3) Using a MRF model, minimize a cost function with the ICM algorithm to obtain an estimate the scale 1 label image. (4) Estimate the scale 0 Gaussian distribution parameters. (5) Estimate the scale 0 label field using the scale 1 label field as an initial estimate. Since we typically expect very different statistical characteristics than those with intensity based lowpass images, we may expect a high-quality segmentation for certain textures. Specifically those with homogeneous regions of different major local edge directions and thus largely, non-overlapping distributions.

In the following we recall the phase of the WPT. We describe the algorithm for MAP estimation and that for the parameter estimation problem. A method for improving phase concentration along major edge directions and thus permitting improved segmentation is described. Finally, two synthetic examples are constructed where good segmentation results are obtained with the new algorithm. These results are compared with the conventional multiresolution algorithm.

2. FEATURE CHARACTERIZATION WITH MULTIRESOLUTION PHASES

Much has been written about the role of phase in image reconstruction and identification. For example, image reconstruction is possible with global Fourier phase [8], local Fourier phase [1] and local Gabor phase [14]. Local Gabor phase has been explored in texture discrimination and image segmentation [11]. There are however, disadvantages when applying this method to real world images. Since textures encountered may vary considerably in frequency and orientation, choosing an appropriate set of Gabor filters is difficult. Long discrete Gabor filters for narrowband characteristic will cover large spatial area in the original image. This deteriorates the locality of the spatial features - a property commonly believed to be important.

In the WPT, with an underlying cyclic group and a quadtree scan of an image, the phase measures the local gradient angle and therefore the local edge direction of the image. This is equivalent to a 4-channel multiresolution Discrete Fourier Transform (DFT) filterbank. For a 2×2 subblock consisting of four elements x_0, x_1, x_2, x_3 scanned clockwise, the phase (of the bandpass coefficient) is:

$$\theta = \arctan \frac{(x_3 - x_1)}{(x_0 - x_2)}. \tag{1}$$

The phase angle represents the angle of the gradient vector and hence the local edge direction.

3. SEGMENTATION WITH MRF MODEL

We consider segmentation within a statistical framework. Specifically, we explore the maximum a posteriori (MAP) estimation of the solution with Bayesian formulation [10]. There are several random processes involved in this problem. We define a neighborhood system first. Let $S = s_1, s_2, \cdots, s_N$ be a set of sites and let $G = G_s, s \in S$ be a neighborhood system for S, which means $s \ni G_s$ and $s \in G_r$ implies pixel $r \in G_s$. A second-order neighborhood system is adopted here, which means that we consider all the eight directly adjacent neighbors. A subset $C \subseteq S$ is a clique if every pair of distinct sites in C are neighbors, and $\mathcal{C}$ denotes the set of cliques. The first random process is the class field containing the classification labels. The label field X is viewed as a locally dependent MRF with respect to G. This means the prior distribution $P(X = x)$ is chosen to be a Gibbs distribution, according to the Hammersley-Clifford theorem. This is expressed as:

$$P(x) = \frac{1}{Z} e^{-U(x)}, \tag{2}$$

where the partition function Z is a normalization constant, and $U(x)$ is the energy function of the following form:

$$U(x) = \sum_{C \in \mathcal{C}} V_C(x). \tag{3}$$

Corresponding to the label field, there is the other random process - the observation process, denoted by Y. An image with intensity or phase values is regarded as a realization of such a process. Assume there are m different classes in the image segmentation problem. Given the class, we assume the observed values are conditionally independent of the others, and have a Gaussian distribution with class dependent parameters $\mu_i, \sigma_i^2, i = 1, 2, \cdots, m$. This is a mixture of Gaussian distributions. Although this assumption is not completely valid, it is reasonable and has been used in [10]. The parameters need to be estimated, which we address later. For the time being, we assume they are all known. Given $P(x)$ and $P(y|x)$, the a posteriori $P(x|y)$ is derived by Bayes theorem:

$$P(x|y) \propto P(y|x)P(x). \tag{4}$$

The MAP estimation of x corresponds to minimization of the cost function:

$$U_e(x) = U_1(x) + U_2(x, y). \tag{5}$$

$U_1(x)$ is a regularization term, and imposes spatial homogeneity constraints into the model. It is defined as follows:

$$U_1(x) = \sum_{C \in \mathcal{C}} V_C(x), \tag{6}$$

where we use for the clique potential:

$$V_c(x_s, x_n) = \begin{cases} -\beta & \text{if} \quad x_s = x_n \\ \beta & \text{if} \quad x_s \neq x_n, \end{cases} \tag{7}$$

where β is assigned some specific value. $U_2(x, y)$ accounts for the link between labels and observations, and can be expressed as

$$\begin{aligned} U_2(x, y) &= \sum_{s \in S} - \ln[p(y_s|x_s)] \\ &= \sum_{j=1}^{N} (\ln \sigma_k + \frac{(y_j - \mu_k)^2}{2\sigma_k^2}), \end{aligned} \tag{8}$$

where k is the class number of x_j. The computational aspects of this nonconvex minimization problem are nontrivial. As described earlier, the ICM algorithm is used for minimization.

4. PARAMETER ESTIMATION

We assume the observation Y to consist of m classes. Each observed value is conditionally independent of the others and has a Gaussian distribution with class dependent parameters $\mu_i, \sigma_i^2, i = 1, 2, \cdots, m$. We consider the maximum-likelihood estimation of these parameters. Let w be the pixel phase, and let Φ denote the parameter vector:

$$\Phi = (\alpha_1, \cdots, \alpha_m, \phi_1, \cdots, \phi_m). \tag{9}$$

Then the probability density function of w depending on Φ can be written as:

$$p(w|\Phi) = \sum_{i=1}^{m} \alpha_i p_i(w|\phi_i), \tag{10}$$

where α_i is nonnegative and $\sum_{i=1}^{m} \alpha_i = 1$, and p_i is a Gaussian function parameterized by ϕ_i:

$$p_i(w|\phi_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-(w-\mu_i)^2/2\sigma_i{}^2}, \quad \phi_i = (\mu_i, \sigma_i{}^2). \quad (11)$$

Suppose $y_j, j = 1, \cdots, N$ is a set of sample observations on the mixture. We are interested in maximizing the likelihood function, i.e., the probability density function of the samples, $p(y|\Phi), y = (y_1, \cdots, y_N)$. For simplicity we assume independent samples. Also for convenience, we consider the log-likelihood function. Following the independence assumption, this can be expressed as

$$L(\Phi) = \sum_{j=1}^{N} \ln p(y_j|\Phi). \quad (12)$$

Maximum Likelihood estimation is that choice of Φ that it maximizes $L(\Phi)$. In the general case, seeking an exact solution to the nonlinear problem is computationally difficult. Here we adopt the Expectation Maximization (EM) algorithm [15], which has many advantages such as reliable convergence, economy of storage and ease of implementation. The EM algorithm is an iterative method, which tries to improve the current estimate with a two-step process at each iteration. If the current estimate is Φ^c, the next estimate Φ^+ will be:

$$\alpha_i^+ = \frac{1}{N} \sum_{k=1}^{N} \frac{\alpha_i^c p_i(y_k|\phi_i^c)}{p(y_k|\Phi^c)}, \quad (13)$$

$$\mu_i^+ = \{\sum_{k=1}^{N} y_k \frac{\alpha_i^c p_i(y_k|\phi_i^c)}{p(y_k|\Phi^c)}\}/(N\alpha_i^+), \quad (14)$$

$$\sigma_i^{2+} = \{\sum_{k=1}^{N} (y_k - \mu_i^+)^2 \frac{\alpha_i^c p_i(y_k|\phi_i^c)}{p(y_k|\Phi^c)}\}/(N\alpha_i^+), \quad (15)$$

for $i = 1, 2, \cdots, m$. The iterative process stops when a predetermined condition is satisfied. For simplicity, we choose m and β (the parameter used in the MAP estimate to specify the linkage strength of a neighborhood) manually.

The EM algorithm is applied in steps 2, and 4 of the segmentation algorithm. The initial values of Φ for the EM algorithm are assigned as follows: First divide the histogram of the scale 1 phase into m parts with equal area. Then take the mean and variance of those values in each part as the initial mean and initial variance. For step 4, the initial value of Φ comes from estimating the coarse label fields, which are obtained from the previous steps.

5. EXPERIMENTAL RESULTS

Some experimental results on segmentation with the WPT phase are presented here. The image in Figure 1(a) consists of a 256×256 image of a wood fence and its $90°$ rotated version. We segment in two parts. The image is transformed first and the bandpass image phase is used to provide the scale 1 segmentation. The result is shown in Figure 1(b) which clearly shows the division into two parts. Figure 1(c) shows the result using the same algorithm as with WPT phase, except the scaled intensity images using lowpass (Haar) filtering is used for the features. Figure 1(d) and 1(e) show the segmentation results after scale 0 refinements.

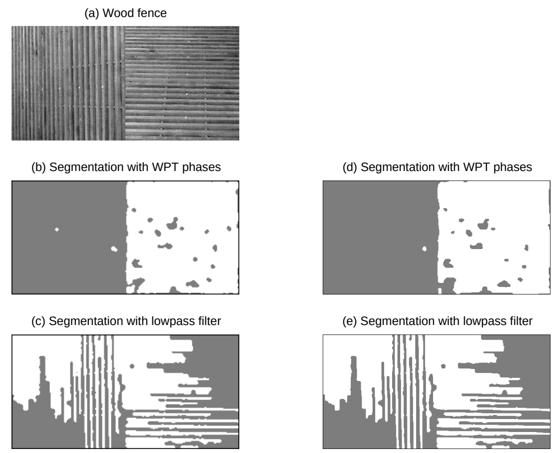

Figure 1: Wood fence segmentation, (b)(c) scale 1, (d)(e) scale 0

Another example is shown in Figure 2. Here, an image of a brick wall is embedded in an image of metal grates. Class number is preset to 2. The WPT phase-based segmentation algorithm is applied, and the scale 1 and scale 0 results are shown in parts (b) and (d). Segmentation results with conventional multiresolution algorithm are displayed in parts (c) and (e), corresponding to scale 1 and scale 0. They conform to gray scale effects, as we expect.

6. DISCUSSION AND CONCLUSION

Modeling phase data with a Gaussian mixture distribution and segmenting an image based on this and a MRF model, leads to segmentation based on local edge directions. To achieve good segmentation, the phase data should have a good concentration along distinct major directions. To achieve this, we use a 5×5 median filter to regularize the phase to within a $[-\frac{\pi}{2}, \frac{\pi}{2}]$ interval. This nonlinear preprocessing greatly improves data concentration, thus improving segmentation. In Figure 3, four images are listed in the first column. They are derived from rotating a simulated texture shown in the upper-left corner, by $0°, 30°, 60°$ and $90°$ degrees respectively. Phase histograms are shown in the second column, and again after median filtering, in the third column. The effect of median filtering in improving data concentration is clear. We observe different phase histograms where intensity histograms would be similar.

By modeling textures as patterns that differ significantly in their dominating edge directions, we have provided a computational multiresolution framework for texture segmentation. We have added local Fourier phase to the classic framework of MRF models with the EM algorithm to predict the Gaussian parameters and iterated conditional modes for image segmentation. Results conform to our visual system for textures having strong local edge directions.

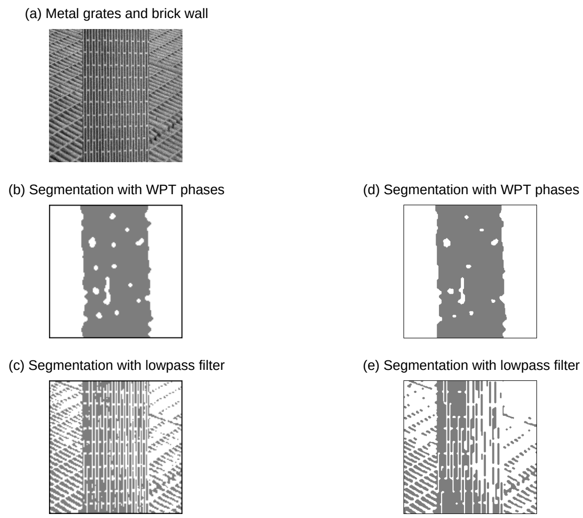

Figure 2: Metal grates and brick wall segmentation, (b)(c) scale 1, (d)(e) scale 0

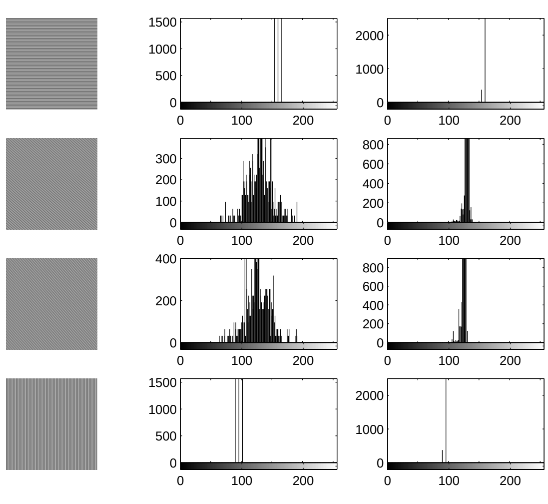

Figure 3: Simulated textures and their phase histograms

7. REFERENCES

[1] J.Behar, M.Porat and Y.Y.Zeeve, "Image reconstruction from localized phase", *IEEE Transactions on Signal Processing*, vol. 40, pp. 736-43, Apr. 1992.

[2] J.Besag, "On the statistical analysis of dirty pictures", *J. Roy. Statist. Soc. B*, vol.48, no.3, pp.259-302, 1986.

[3] J.Besag, "Spatial interaction and the statistical analysis of lattice systems", *J. Roy. Statist. Soc. B*, vol.36, no.2, pp.192-236, 1974.

[4] S.Geman and D.Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images", IEEE Trans. PAMI, vol.6, no.11, pp.721-741, 1984.

[5] E.G.Huot and I.L.Herlin "Cropland detection with SAR interferometry: A segmentation model", *Proceedings of ICASSP*, vol 5, pp.2733-2736, 1998.

[6] F.Luthon, A.Caplier and M.Lievin, "Spatiotemporal MRF approach to video segmentation: application to motion detection and lip segmentation", *Signal Processing*, vol.76, pp.61-80, 1999.

[7] G.Mirchandani, R.Foote, D.Rockmore, D.Healy and T.Olson, "A wreath product group approach to signal and image processing: part II-convolution, correlation, and applications", *IEEE Transactions on Signal Processing*, March, 2000, vol.48, No.3, pp.749-767.

[8] A.V.Oppenheim and J.S.Lim, "The importance of Phase in Signals", *Proceedings of the IEEE*, vol.69, No.5, pp.529-541, May, 1981.

[9] N.Paragios and G.Tziritas, "Adaptive detection and localization of moving objects in image sequences", *Signal Processing: Image Communication*, vol.14, pp.277-296, 1999.

[10] G.Poggi and R.P.Ragozini, "Image segmentation by tree-structured Markov random fields", *IEEE Signal Processing Letters*, vol.6, no.7, pp.155-157, 1999.

[11] M.Porat and Y. Zeevi, "Localized texture processing in vision: Analysis and synthesis in the Gaborian space", *IEEE Trans. Biomedical Engineering*, vol.36, No.1, pp.115-129, Jan. 1999.

[12] M.Saeed, W.C.Karl, T.Q.Nguyen and H.R.Rabiee, "A new multiresolution algorithm for image segmentation" *Proceedings of ICASSP*,vol.5,pp.2753-2756, 1998.

[13] L. Wojnar, *Image analysis: Applications in Materials Engineering.* CRC Press, 1999, pp.170-171.

[14] Y.Y.Zeevi and M.Porat, "Computer image generation using elementary functions matched to human vision," *Theoretical Foundation of Computer Graphics*, New York: Springer, 1988, pp. 1197-1241.

[15] J.Zhang, J.Modestino and D.Langan, "Maximum likelihood parameter estimation for unsupervised stochastic model-based image segmentation", *IEEE Trans. Image Process.*, vol.3, no.4, pp.404-420, 1994.